Semantic opinion mining of social media for stock market analysis

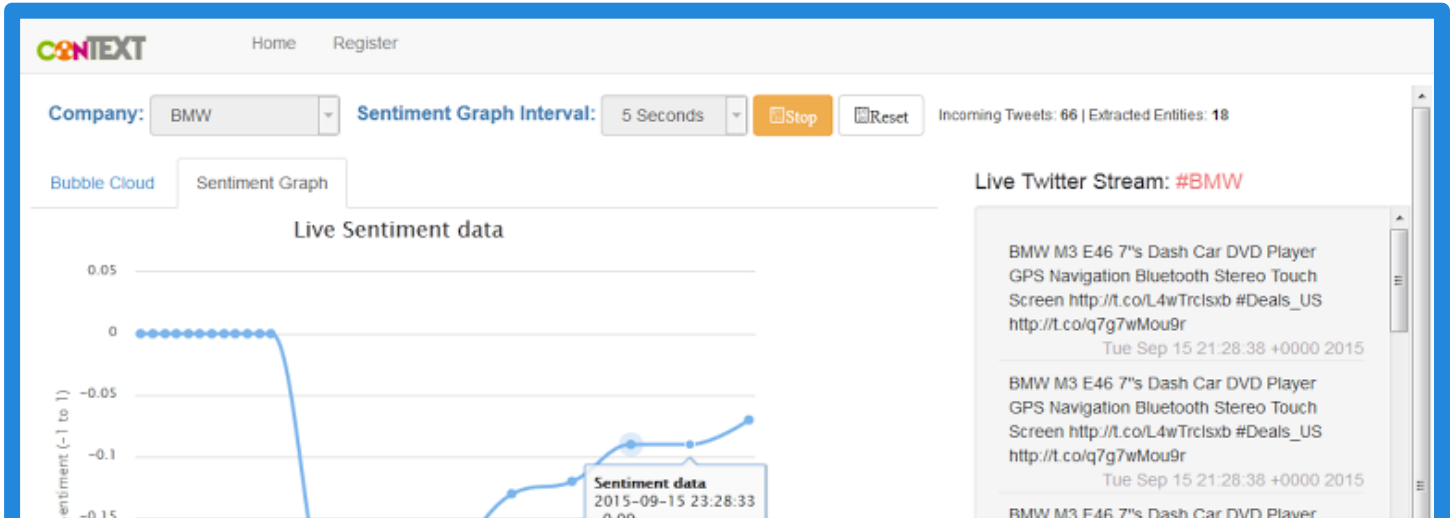
Introduction

Presented by
Cristobal Leiva
Supervised by
Dr. Simon Scerri
Prof. Dr. Sören Auer

universitätbonn EIS ENTERPRISE INFORMATION SYSTEMS

Previous work...

- Tracking stock market sentiment through the semantic analysis of social media
  - Author: Priyanka Dank - 2014
  - github.com/EIS-Bonn/Theses/tree/master/2014/Priyanka_Dank



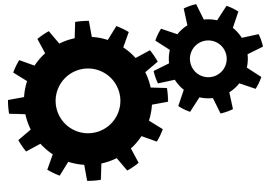
Motivation

- Opinion of the public can be extracted from the social media
- People's opinion towards a company may influence value changes in the stock market
- Stock market movements can be predicted by analyzing the emotional behavior of the people using social media content

Problem

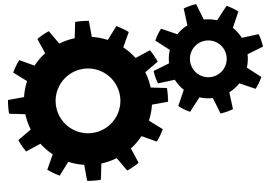
- Sentiment classifiers problems
  - Heterogeneity of posted information
  - Message based analysis
- Sentiment expressed towards an entity such as a Company might not be enough to represent its global opinion
  - Constant release of new products
  - Change of key personnel

Approach/Solution



- Social networks to be analized:
  - **Twitter** and Facebook (in evaluation)
- Entity-based sentiment analysis using semantic web technologies
  - Build an **ontology** to represent companies entities
  - Extraction of Company-specific data from different sources such as DBpedia, Google finance and others.
- Extend ReSA (Real time Semantic Analysis tool by Dr. Ali Khalili)
  - Extend DBpedia spotlight data
  - Process streams from different social networks
  - Replace sentiment analysis module

Approach/Solution - (Contd.)



- 3-class classification: Negative / Positive / Neutral
- Entity-base sentiment classifier
  - Provide sentiment towards given entity (-1.0 /1.0)
- Combined Techniques
  - Combine supervised and unsupervised classification methods.
- Lexical resources
  - NRC Emotion (Twitter specific tokens)
  - MPQA (8,000 words)
  - SentiWordNet (3.0)

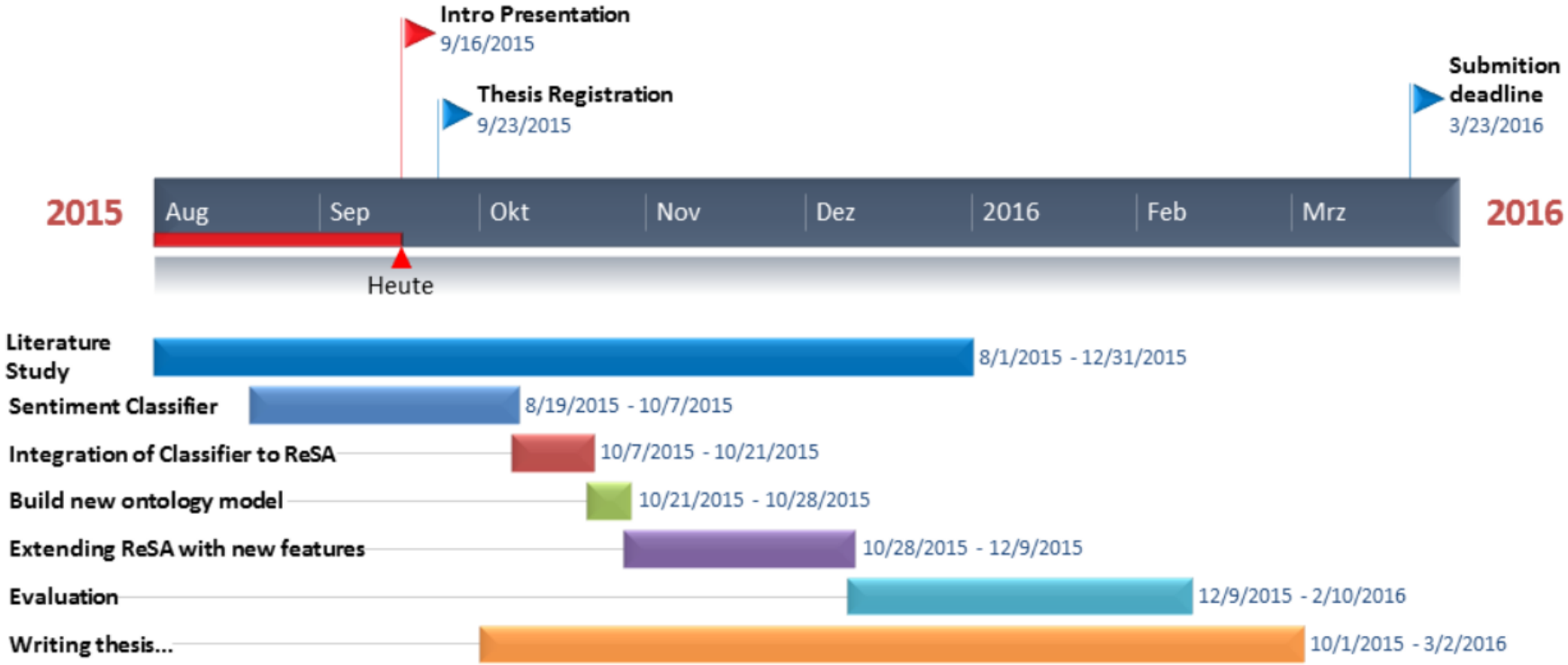
Implementation - Technologies

- Extending ReSA project
  - NodeJS / MongoDB / Twitter API
  - + FacebookAPI / Several Lexicon Libs
- Entity-expansion based on created ontology
  - RDF / SPARQL
  - DBpedia / DBpedia Spotlight

Evaluation

- Sentiment classifier evaluation
  - Annotated evaluation datasets (SemEval competition)
- Find correlations between average sentiment and stock market trend for given entity
  - Independent application recording stock value changes

Time plan


Intro Presentation
9/16/2015
Thesis Registration
9/23/2015
Submition deadline
3/23/2016
2015
Aug
Sep
Okt
Nov
Dez
2016
Feb
Mrz
2016
Heute
Literature Study
8/1/2015 - 12/31/2015
Sentiment Classifier
8/19/2015 - 10/7/2015
Integration of Classifier to ReSA
10/7/2015 - 10/21/2015
Build new ontology model
10/21/2015 - 10/28/2015
Extending ReSA with new features
10/28/2015 - 12/9/2015
Evaluation
12/9/2015 - 2/10/2016
Writing thesis...
10/1/2015 - 3/2/2016

Current status

- Entity-base Sentiment Classifier for Twitter
  - Selection of lexicons to use
  - Classify several entities on a same tweet

- Evaluating different approaches
  - Other supervised classifiers
  - Usage of services such as Sentiment140 or AlchemyAPI

Thank You

Presented by
Cristobal Leiva

Supervised by
Dr. Simon Scerri
Prof. Dr. Sören Auer